4

Studies of Neck Injury Criteria Based On Existing Biomechanical Test Data

G. S. Nusholtz, L. Di Domenico, Y. Shi, P. J Eagle, Z. Ben Aoun

ABSTRACT

This paper presents two analytical approaches to evaluate the Nij injury criterion using previously published biomechanical data. A direct fit optimization analysis is applied to re-examine the intercepts used in the Nij criterion. The results show that no values of the intercepts used in the Nij criterion corresponding to compressive force, extension moment and flexion moment will optimally separate the injury and non-injury cases in the biomechanical data set. Only the tensile intercept allows identification of the injury risk. The tension and bending moment data in the same test series were examined for correlation with injury prediction using a variety of standard statistical techniques. All indications are that increasing the weight given to moment decreases the accuracy of the prediction. Moment is a confounding random variable in these biomechanical data sets. Only neck tension is shown as a reliable predictor of injury potential.

INTRODUCTION

Determining realistic biomechanical injury criteria for impact conditions is a formidable task. The results and conclusions from biomechanical impact studies are incomplete and potentially invalid. This is a consequence of the inability to validate the results, the multiple mechanisms for a given injury, the non-uniqueness of an injury for a given input as well as the complexity of human anatomy, material properties and structure. Only a limited number of tests can be conducted due to cost and complexity. Either postmortem-human or animal human surrogates must be used. In addition, there is never access to the real physical phenomena of an impact event. Tests are always idealized and are limited by experimental errors, assumptions and constraints. Finally, the tool used to estimate the parameters of human response, the anthropomorphic test device (dummy), may not represent the response of a living subject. Therefore, the results from any procedure used to determine assessment values, risk of injury or injury threshold might be distorted or potentially erroneous. However, the need remains for estimating human performance in impact. Efforts should be focused on improving or correcting current assessment values and threshold values.

The neck assessment values have recently been an active area of investigation. This is primarily the result of the re-examination of the live animal test data that was used to establish assessment values in the Federal Motor Vehicle Safety Standard, FMVSS 208, the regulation for frontal crash occupant safety protection.

The human neck is a complex biomechanical structure that can sustain loading due to axial forces and bending moments acting on the spinal column. The mechanisms of injury are related to these loads. Automotive crash events impart time-varying loads on the human neck. Researchers have tried to characterize the injury potential of these complex loading conditions in the form of injury criteria. An injury criterion that combines the effects of tension and moment, N_{ij}, has been proposed. Current regulations governing the performance of occupant safety protection systems have generally been based on the research used to formulate the N_{ij} criterion.

The objective of the work is to determine if the existing criteria are rigorously supported by the test data as good indicators of neck injury risk. This paper presents two analyses that use the same biomechanical data that is the basis for N_{ij}. The first analysis is a direct fit optimization that re-examines the constants used in the N_{ij} criterion. The second analysis follows a statistical methodology and examines whether tensile force and extension bending moment make a significant contribution to the prediction of injury. The appropriateness of neck injury criteria that use combinations of tension and bending moment is assessed. This study also suggests what type of criterion the available test data support.

BACKGROUND

Neck injury criteria are used to assess whether protection of the occupant's neck complies with a regulation. The neck injury criterion, like other injury criteria, is derived from a biomechanical test database using some type of data analysis methodology. Neck injury due to inertial load of the head while the torso is restrained by a seat belt or seat back has been the subject of some study. However, the biomechanical test database for the more complex loading condition resulting from air bag interaction consists primarily of a series of tests conducted to study the injury potential to an out-of-position occupant. Mertz, et al., (1982b), and Prasad and Daniel, (1984), reported paired sets of matched tests conducted on piglets and a crash dummy representing 3-year old children. Small pigs were chosen to represent the size, weight and state of tissue development of three-year-old children. The same conditions (speed, air bag type and subject position) were used on both the piglet and the dummy.

Broad ranges of test conditions were involved, resulting in a spectrum of injuries to different body regions of the piglets. The level of injury resulting from this particular condition is determined by necropsy of the animal and the mechanical response of acceleration and forces are recorded from the dummy. Analysis of the test data is given in (Mertz et al, 1982b, Prasad and Daniel, 1984). Further analyses are given by Mertz and Weber, (1982a), and Mertz et al., (1997).

Mertz and Weber, (1982a), examined the relationship between neck injury level and the peak neck tension force. With a statistical procedure, a neck injury risk curve was established, which shows the percentage of porcine subjects with significant neck injuries as a function of the peak neck tension. The results showed that with a peak tension force of about 1050 N to 1350 N, the risk of injury increases from near zero to near 100%.

Prasad and Daniel, (1984), analyzed the dependence of neck injury severity on the loading represented by three different measures: peak neck tensile force, peak neck extensional moment and a specific combination of the two. Their test data, which were derived from the same test setup and procedure as the Mertz and Weber tests, indicated that all piglets sustained significant neck injuries for peak tensile forces above 1925 N. When the dependence of the injury on the peak extensional moment was examined, the authors suggested that injury severity depended on a combination of tensile force and moment. The injury was assumed to depend linearly on combined peaks of tensile force and extensional moment at a given time. It was concluded that a straight line, passing through the tensile force axis at 2000 N, and the extensional moment at 34 Nm, appears to delineate the no neck injury zone and the severe injury zone.

Mertz et al., (1997), combined and analyzed the data reported in (Mertz et al., 1982b) and (Prasad and Daniel, 1984) using the statistical method reported in (Mertz and Weber, 1982a). The results of this analysis for tension alone create a one- percent risk of AIS $\geq$ 3 injury at 1070 N. When the peak extensional moment alone is considered, 13.0 Nm corresponds to the same injury risk. The injury is related directly to the peak value of a specific combination of the force and moment, N_{TE}, which uses a simple cross-neck force distribution model:

$$(1) \qquad N_{TE} = f(M_E + \frac{D}{2}F_T)$$

where M_E and F_T are the extensional moment and tensile force as functions of time, D is a constant related to the geometry of the neck and f is a constant that is determined from injury and force and moment data. Based on the test data in (Mertz et al., 1982b) and (Prasad and Daniel, 1984), the coefficient f can be determined. This results in the following:

$$(1') \qquad N_{TE} = \frac{M_E}{M_C} + \frac{F_T}{F_C}$$

where M_c is 20.0 Nm and F_c is 1590 N for 3-year olds.

Neck injury assessment values first became part the FMVSS 208 regulation in 1997. Individual limits are given for forces and moments of the dummy neck during a sled test. A new neck injury criterion, Nij, was introduced in September 1998 in a Notice of Proposed Rulemaking (NPRM) for FMVSS 208 (NHTSA, 1998). Nij is defined as:

$$(2) \qquad N_{ij} = \frac{F_z}{F_{zc}} + \frac{M_y}{M_{yc}}$$

where F_z and M_y are the axial force and the neck flexion-extension moment as functions of time. F_{zc} and M_{yc} are normalization constants. A value F_{zc} is given when F_z is tensile force and a possibly different F_{zc} value is used when F_z is a compressive force. Similarly, two values are used for M_{yc} depending on whether M_y is extension or flexion. F_{zc} and M_{yc} values are dependent on the representative gender and age of the dummies (partly based on the biomechanical data reported in (Mertz et al., 1982b, Prasad and Daniel , 1984, Mertz and Weber, 1982a, Mertz et al., 1997)). These values were modified in a Supplemental Notice of Proposed Rulemaking (SNPRM) for FMVSS 208 (NHTSA, 2000b). They were further revised in 2000 in the FMVSS 208 Final Rule (NHTSA, 2000a). In this regulation, the values of F_{zc} and M_{yc} for the 3-year old child dummy are as follows:

$$Fzc = 2120 \text{ N when Fz is in tension,}$$
$$Fzc = 2120 \text{ N when Fz is in compression,}$$
$$Myc = 68 \text{ Nm when My is flexion and}$$
$$Myc = 27 \text{ Nm when My is in extension.}$$

Four values of N_{ij} are computed as functions of time. The four values correspond to the possible combinations of axial force and moment: tension-extension, tension-flexion, compression-extension and compression-flexion. In addition to limiting any of the Nij values to 1.0, the regulation further requires that the peak tension force shall not exceed 1130 N and the peak compression force shall not exceed 1380 N at any time for the 3-year old child dummy. F_{zc} and M_{yc} values for dummies of other sizes are also given and they are essentially arrived at by scaling the above numbers.

The regulatory requirements for the 3-year old dummy can be depicted in M_y - F_z space as shown in Figure 1. The N_{ij} requirement is represented as a diamond shape. In this depiction, the values for M_{yc} and F_{zc} correspond to the intercepts of the N_{ij} limit lines on the axes. The peak tension and compression limits appear as dotted vertical lines.
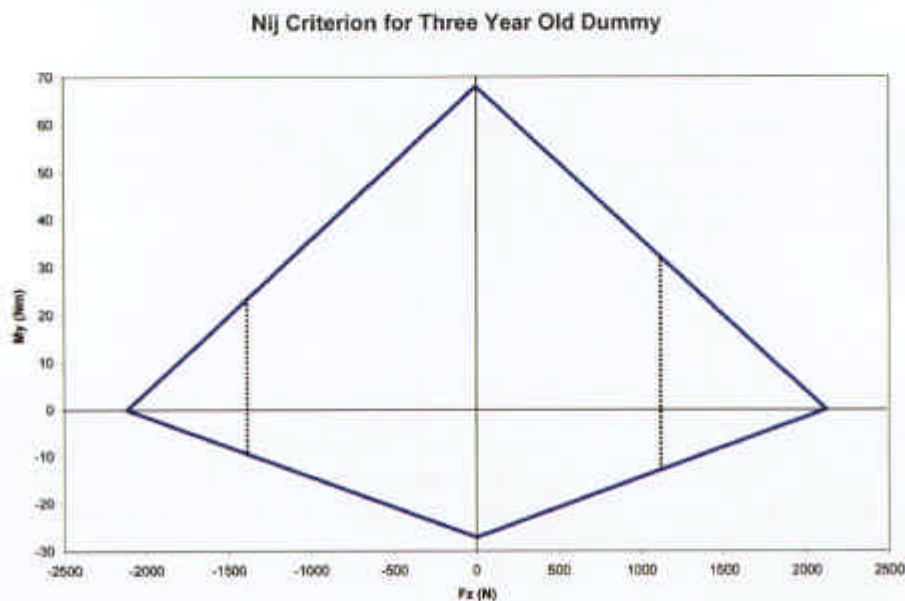
Nij Criterion for Three Year Old Dummy

Figure 1. Neck injury protection requirement envelope of FMVSS 208 regulation for a 3-year-old test dummy.

EXISTING BIOMECHANICAL DATA USED IN ANALYSIS

Out of the forty-six tests (forty-three piglet plus three baboon tests) reported in (Mertz et al., 1982b), there was enough usable data in thirty-three tests. Twelve tests out of the fifteen reported in (Prasad and Daniel, 1984) contained enough data to be included in this study. Therefore, forty-five paired piglet-dummy tests were used in this study. No distinction is made between those from (Mertz et al., 1982b) and those from (Prasad and Daniel, 1984) because the test methodology and setup were such that the tests can be considered as a single test program.

Necropsies of the pigs allowed injury severities to be correlated with corresponding measurements made on the test dummy. The dummy neck force and moment time histories and the corresponding injury level from the piglet data from a subset of the tests that are reported in (Mertz et al., 1982b, Prasad and Daniel, 1984) were obtained from their respective original sources (Mertz, 1999, Kim, 1999) for this study.

Only the part of the time histories that are considered relevant to the air bag loading should be used for the analysis. Efforts were made to assure that the selection of the time duration was consistent with that used in (Mertz et al., 1982b, Prasad and Daniel, 1984, Mertz and Weber, 1982a, Mertz et al., 1997).

DIRECT FIT ANALYSIS OF THE NIJ CRITERION

This analysis assumes that neck injury is governed by the N_{ij} as defined in Equation (2) where F_{zc} has unique values for tension and compression and M_{yc} has unique values for flexion and extension. The objective is to use the existing test data described above to determine four intercept values that allow a best discrimination between injury and non-injury cases. Because of the nature of the tests and types of injuries observed in the subjects, flexion and compression should not be explainable values and should be not related to injury. They are included as a test of the procedure. If the procedure is valid, there should not be a value for the flexion and compression intercept that offers a clear advantage in injury prediction.

This approach seeks to minimize, with respect to $\overline{L}$ (a vector of all four to-be-determined intercepts), a fit measure p that allows the "best" separation of injury and non-injury tests, per some requirement (a similar method is presented in (Nusholtz et al., 2000)). Figure 2a and 2b show the time histories of the tests in the tension-moment space along with the diamond-shaped failure envelope corresponding to $N_{ij} = 1$.
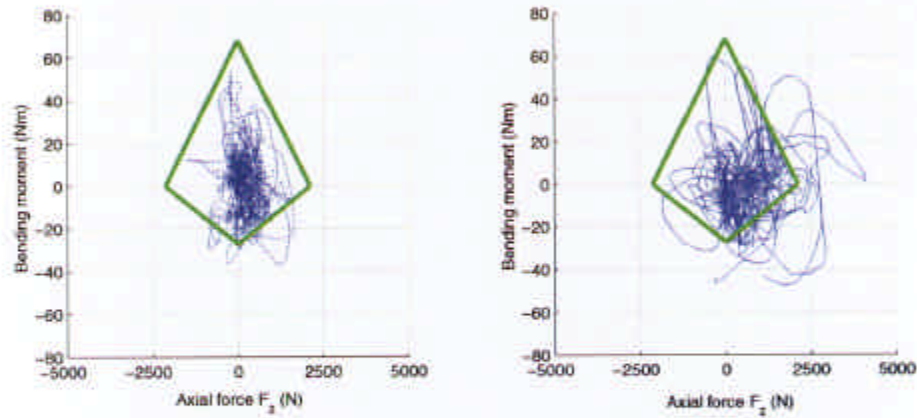
Figure 2. Time histories of no-to-low injury cases (a) and the serious-to-fatal injury cases (b).

The fit measure p is defined as:

$$(3) \qquad p(\overline{L}) = \left(\sum_{i=1}^{45} (\alpha_i (cl_i, I_i) cl_i - I_i)^2 \right)^{1/2}$$

where cl_i is a "converted" injury classification for the ith test based on the N_{ij} value: $cl_i = 1$ ("serious-to-fatal injury"), if $N_{ij} \geq 1$ and $cl_i = 0$ ("no-to-low injury") if $N_{ij} < 1$. I_i is the injury classification from the ith test result: $I_i = 1$ if $AIS_i \geq 2.5$ and $I_i = 0$ if $AIS_i < 2.5$. α_i is a weighting factor given to the ith test that allows the fit to be carried out for different requirements.

The similarity between the N_{ij} criterion and the tension-extension based criterion reported in (Mertz et al., 1997) is apparent. Further, the determination of the values of M_c and F_c in (Mertz et al., 1997) and the determination of the intercepts through the fit analysis in this study appear to be similar. However, two differences are worth noting here. Most importantly, in (Mertz et al., 1997), there is a preset relationship between Mc and Fc, as given by Equations (1) and (1'), which is a result of a simple cross-neck force distribution model assumption. Secondly, only the tension and extension intercepts are considered as in (Mertz et al., 1997). In this study, all of the intercepts (tension and extension as well as compression and flexion) are considered.

Two different fit requirements were examined:

(1) $\alpha_i = 1$. The same weight is given to all cases; therefore, the final fit given by minimizing p is one that makes the minimal number of "incorrect" predictions of injury by the N_{ij} criterion.

(2) $\alpha_i = w$, if $cl_i = 0$ and $I_i = 1$; else $\alpha_i = 1$. The weighting factor, w, is a large number given as a penalty to the case where the actual test is an injury condition, but the N_{ij} predicts non-injury. The incorrect prediction of an actual injury case is penalized more than the incorrect prediction of an actual non-injury case. Therefore, the final fit given by minimizing p is one that predicts all actual injury cases correctly and, at the same time, makes the minimal number of incorrect predictions about the actual non-injury cases with a suitably selected value of w (in this study, w = 25 was used).

The direct fit analysis was applied to both the case where all four intercepts were allowed to vary and the case where only the tension-extension intercepts were allowed to vary. The latter case assumes that the observed injury was all from this mode of loading.

The requirement for a minimal number of incorrect predictions (i.e., the fit when $\alpha_j = 1$) was considered with all four intercepts allowed to vary. The response of p is represented by the mesh response surface in Figure 3. The response surface is largely a cylindrical surface after the extension moment intercept exceeds about −70 Nm − indicating that only one variable, tension, affects the injury predominantly. There is a value for the tension intercept, approximately 1225 N, which minimizes the number of incorrect predictions. The limits for compression, extension and flexion do not appear to significantly affect the minimization of incorrect injury predictions. Figures 5, 6 and 7 show that the fit measure, p, cannot be minimized by selecting any value of the extension, compression or flexion limits in the range of the existing data. The fit measure is not effected by increasing or decreasing the extension, compression or flexion limits beyond some value in the range of the available data. This suggests that fixing these parameters to a specific value will not contribute to accurately predicting injury in the range of available test data.
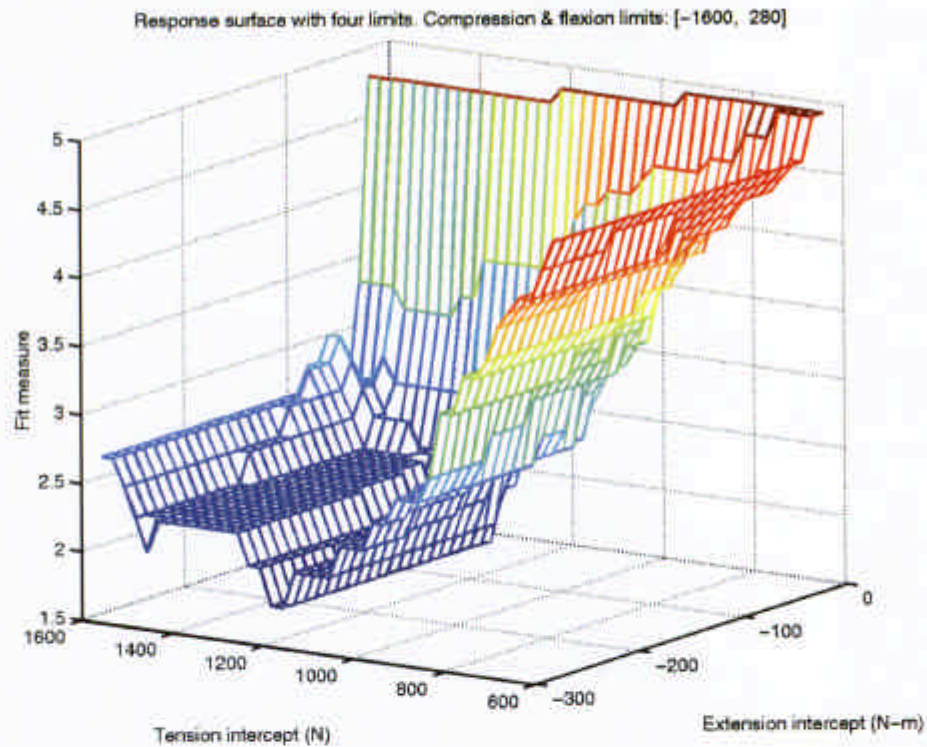

Figure 3. Response surface of the fit measure with the first requirement and four intercepts allowed to vary.
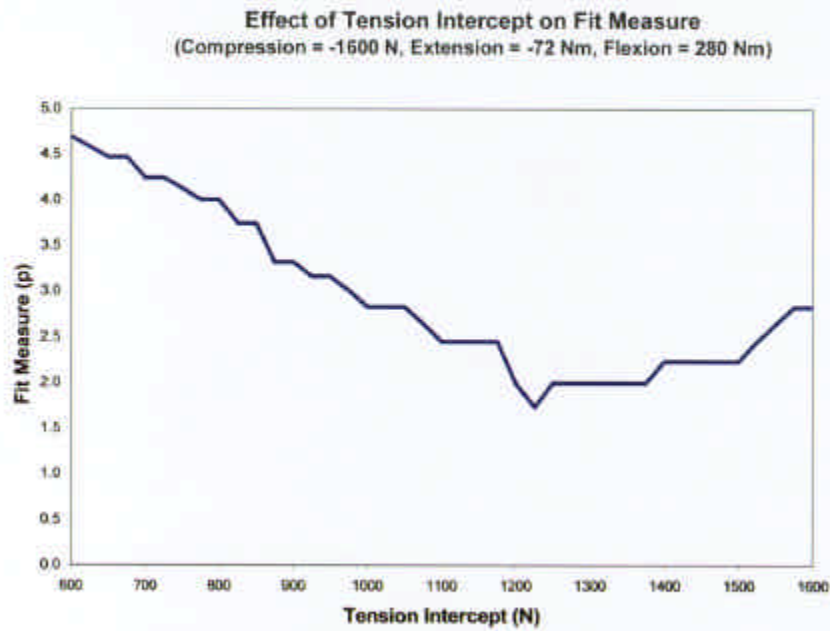
Effect of Tension Intercept on Fit Measure
(Compression = -1600 N, Extension = -72 Nm, Flexion = 280 Nm)

Figure 4. Sensitivity of fit measure to the tension intercept with the first requirement when four intercepts are allowed to vary.

Effect of Extension Intercept on Fit Measure
(Tension = 1225, Compression = -1600 N, Flexion = 280 Nm)
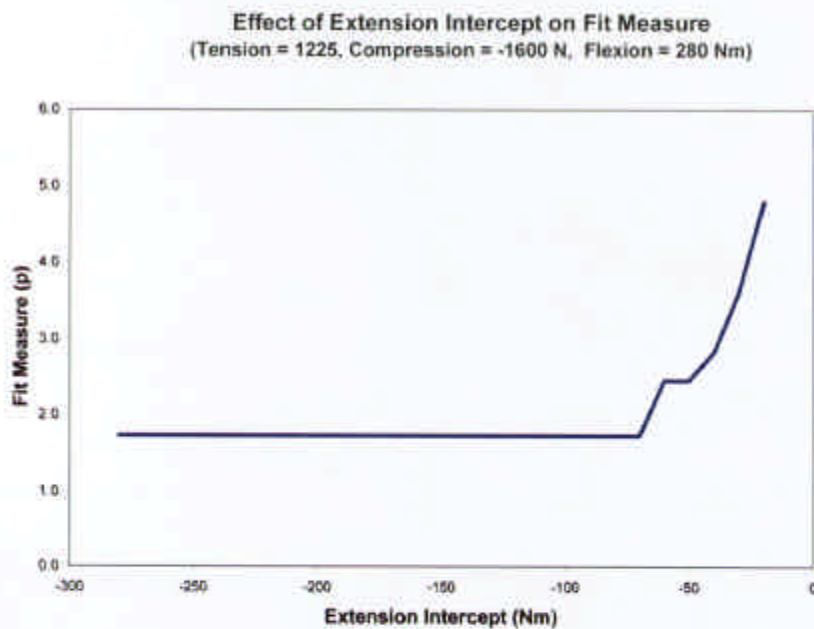

Figure 5. Sensitivity of fit measure to the extension intercept with the first requirement when four intercepts are allowed to vary.

Effect of Compression Intercept on Fit Measure
(Extension = -72 Nm, Flexion = 280 Nm, Tension = 1225 N)

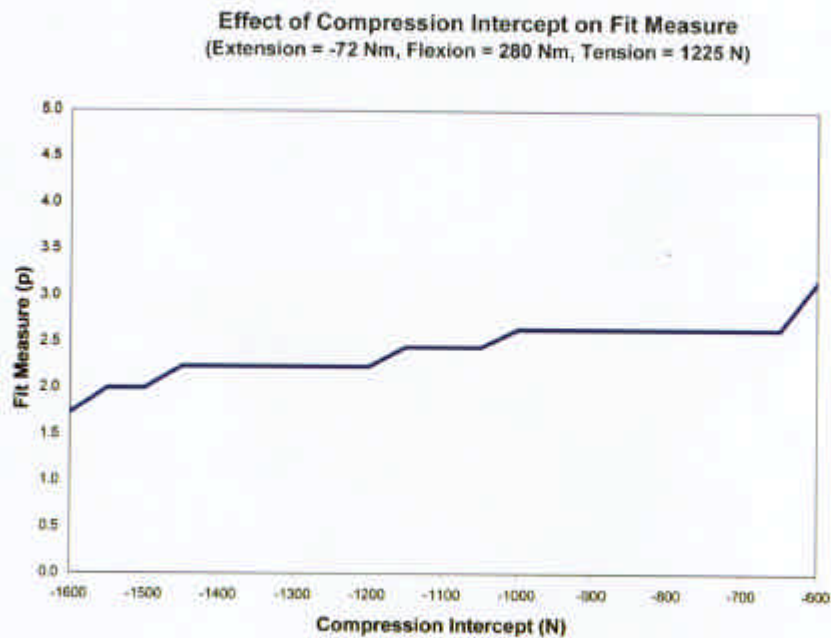Figure 6. Sensitivity of fit measure to the compression intercept with the first requirement when four intercepts are allowed to vary.

Effect of Flexion Intercept on Fit Measure
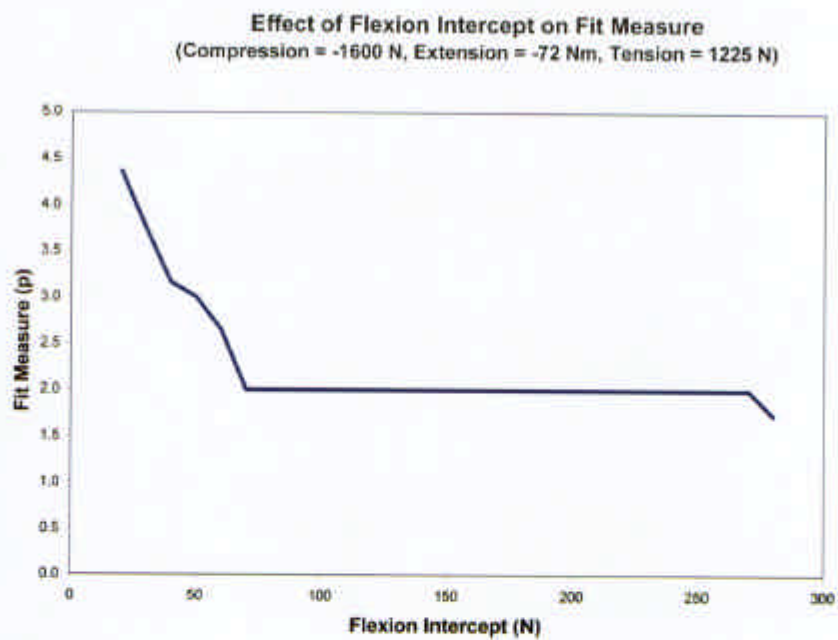(Compression = -1600 N, Extension = -72 Nm, Tension = 1225 N)

Figure 7. Sensitivity of fit measure to the flexion intercept with the first requirement when four intercepts are allowed to vary.

The requirement for a minimal number of incorrect predictions was also considered with only the tension and extension intercepts allowed to vary and only the tension and extension parts of the data are used. The response of p to different values of the two limits is given in the mesh response

surface shown in Figure 8. The response is very similar to the four-parameter case. The fit measure, p, is minimized by a single value of the tension intercept. Specifically, the tension intercept is clearly defined, while the extension intercept does not affect the quality of the fit once it is beyond a certain value. Figure 9 shows that p cannot be minimized by selecting any value of the extension in the range of the existing data.
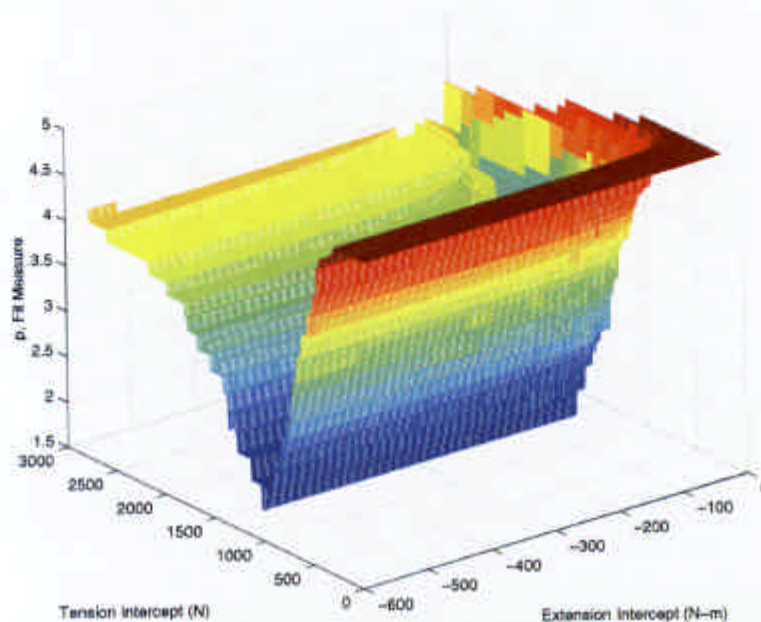

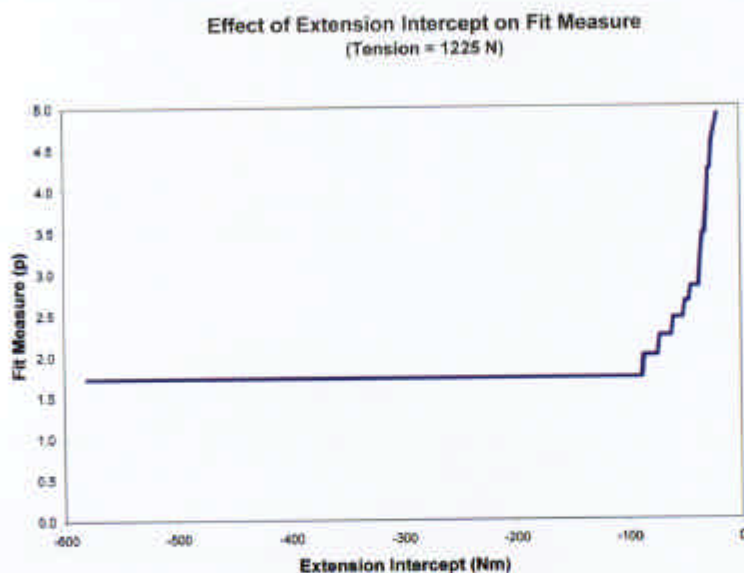Figure 8. Response surface of the fit measure with first requirement and two intercepts allowed to vary.

Figure 9. Sensitivity of fit measure to the extension intercept with the first requirement when two intercepts are allowed to vary.

In both the four-parameter and two-parameter cases, the tension limit is the only parameter that effectively minimizes the number of incorrect predictions of injury. The range of the existing data does not appear to support the use of compression, extension or flexion in attempting to accurately predict injury.

The same type of investigation was conducted for the case where minimizing requires predicting all actual injury cases correctly and making the minimal number of incorrect predictions about the actual non-injury cases (i.e., the fit when $\alpha_i = w$). The results are shown for the four-intercept case in Figure 10. There is a value for the tension intercept (again, approximately 1225 N) which minimizes the fit measure.
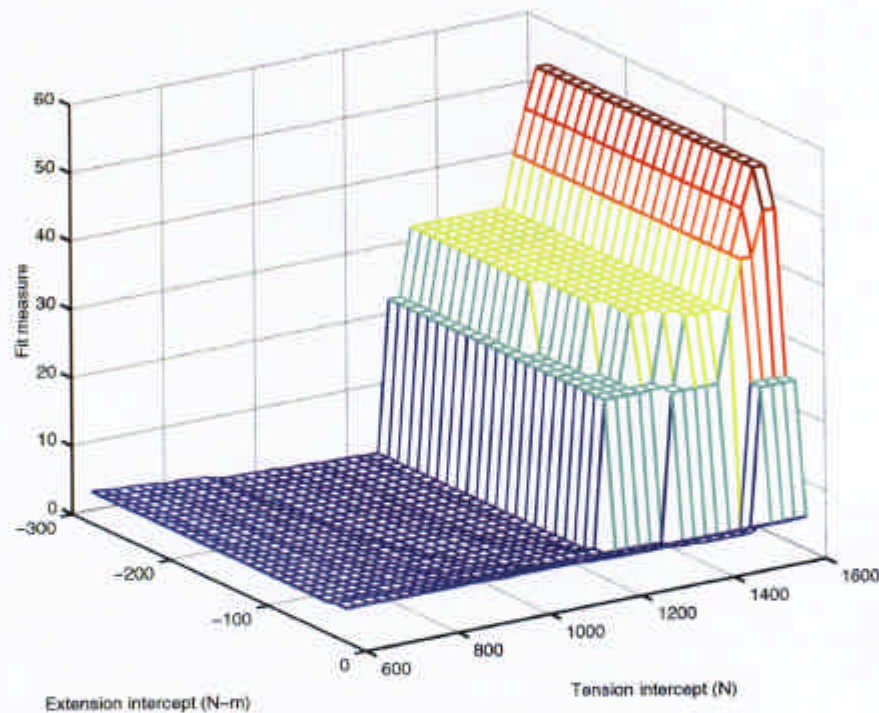

Figure 10: Response surface of the fit measure with the second requirement and four intercepts allowed to vary.

This requirement was also examined for the two-intercept case. The results are shown in Figure 11. There is a value for the tension limit that minimizes the fit measure (approximately 1160 N). No value of extension in the range of available test data minimizes the fit measure. There is a small difference between the intercept values for tension obtained in the four-intercept and two-intercept cases. This is due to the trajectories of some no-to-low injury time histories passing through the tension-flexion quadrant as well as differences in step sizes in the search process.
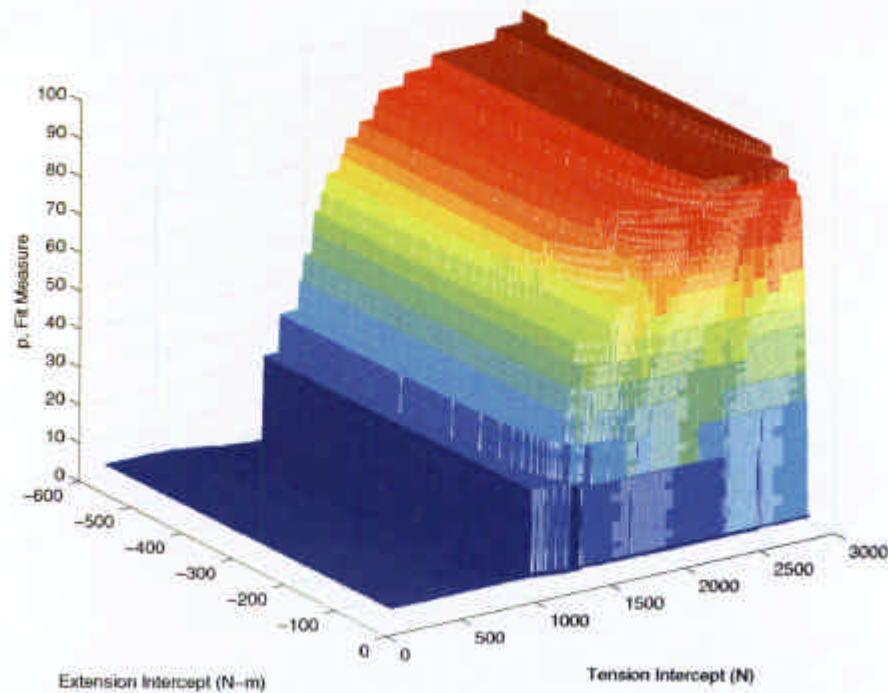
Figure 11: Response surface of the fit measure with the second requirement and two intercepts allowed to vary.

Table 1 summarizes the optimal intercept values found by the direct fit optimization method. The "±" symbol indicates the step size in calculating the response, and therefore, the uncertainty in the value. The "≤" and "≥" means the response (i.e., the prediction of the Nij with these intercepts of the test results) remains the same below or above this value. The minimum value of the fit measure is not influenced by extension, compression and flexion in the range of available data. Insensitivity of the prediction to the intercept values is exhibited, except in the case of tension.

Table 1. SUMMARY OF INTERCEPT VALUES.

Intercepts	Tension Intercept	Extension Intercept	Compression Intercept	Flexion Intercept
Four Intercepts	1225 ± 25 N	$\leq -70 \pm 10$ Nm	≤ -1600 N	≥ 280 Nm
Two Intercepts	1159 ± 9 N	$\leq -88 \pm 2$ Nm	N/A	N/A

It is noted that the optimal points for the four-intercept cases for these two requirements (both minimizing the number of incorrect predictions and prohibiting incorrect actual injury prediction with fewest incorrect non-injury predictions) are identical. The same is true for the two-intercept cases for both requirements. This can only be true when the incorrect predictions for the first type of requirement are all about the non-injury test cases. When the actual test results and the predictions for the four-intercept and two-intercept cases are compared, the same 3 of 45 tests are miss-classified. The use of the four intercepts does not improve the predictions of injury with this particular set of test data.

STATISTICAL ANALYSIS OF TENSION AND MOMENT

The results of the direct fit analysis calls into question whether the use of moment in a neck injury criterion contributes to the accuracy of injury prediction for air bag loading. To investigate further, two different analyses, principal component and confidence tests on the mean, were performed to investigate the predictive value of tension and extension moment. The direct fit analysis indicated that tension was the significant variable. For this reason, tension alone was considered. The focus was on determining whether extension moment makes a significant contribution to predicting a neck injury beyond what is seen with tension. The statistical analysis presented here shows that there is evidence to conclude that moment cannot provide additional separation between injury and non-injury in the experimental data set beyond the separation already achieved by the tension alone. The same data used in the direct fit optimization study was used in these analyses. The global peak values of tension and moment were used in this analysis as compared to the global peak values used in the calculation of N_{ij}. The individual variables will exhibit a stronger influence on injury using this approach. If moment is not found to be significant when its global peak is considered, it is unlikely to be a significant contributor to injury when it is applied in the N_{ij} criterion.

Preliminary Analyses

Preliminary analyses (tests on the mean value) were performed one for each of the two sets of data (Mertz et al., 1982b) and (Prasad and Daniel, 1984) and one for the combined data set. The preliminary analysis for the data in (Mertz et al., 1982b) reached the same conclusion suggested by the analysis of the combined data with the same power. On the other side it is not possible, using only the data in (Prasad and Daniel, 1984), to draw any reliable conclusion: the data set is too small to insure enough power and robustness of the analysis. All the statistical analyses described in what follows were performed using the combined data reported in (Mertz et al., 1982b, Prasad and Daniel, 1984).

Principal Component Analysis

Principal component analysis seeks to maximize the variance of a linear combination of the variables under consideration (the peak values of tension (T) and moment (M)). The Abbreviated Injury Scale (AIS) values (i.e., designation of serious-to-fatal or no-to-low injury) will not be used in this analysis. All samples are combined together and no grouping of observations are assumed. The principal component analysis is concerned only with explaining the variance-covariance structure of the tension and moment variables through a linear combination that has maximal variance. Seeking a linear combination with maximal variance is essentially searching for a line that the observations can be projected onto that creates the largest separation among the observations.

The first principal component is given by (Johnson and Wichern, 1998):

$$(4) \qquad P_1 = 1.000\,(T - E(T)) + 0.005\,(M - E(M))$$

where $E(T)$ and $E(M)$ are the expected or mean values of the tension and moment where T is measured in Newtons and M is measured in Newton-meters. The relative sizes of the coefficients of tension and moment suggest that tension contributes significantly more to the determination of P_1. The test data do not show significant separation in terms of moment. However, tension, taken essentially alone, separates the data as widely as possible. This suggests the possibility that if the data is grouped into serious-to-fatal injury and no-to-low injury categories. Tension may be used to discriminate between the two categories.

The correlation between the first principal component and the two variables was also examined. The correlation values were found to be:

$$Corr\ (P_1, T) = 1.000 \quad and \quad Corr\ (P_1, M) = 0.004$$

The first principal component is perfectly correlated with tension and, for all practical purposes, not correlated with moment. This implies that tension determines the value of the principal component and, hence, the scatter that exists in the data. The first principal component explains 99 % of the total variance. Therefore, P_1 can "replace" the original variables without practical loss of information.

The principal component analysis provides evidence to conclude that tension is the most important variable in this data set while moment appears to not be significantly informative. The results suggest that the principal component essentially duplicates the tension variable. Tension alone is able to explain 99% of the total information contained in the data. Thus, the principal component analysis gives evidence to suggest that moment can be omitted from the analysis without loss of information.

Tests on Mean Values

Tests were also conducted on whether the means values of tension and moment could be used to separate the data into serious-to-fatal injury and no-to-low injury categories. In this analysis, the AIS variables were used to categorize all the available observations into Group 1 (serious-to-fatal injury) and Group 0 (no-to-low injury). Group 1 is the family of all observations where AIS ≥ 3. Group 0 is the family of all observations where AIS < 3. This test examines whether tension (T) and moment (M) can be used to separate the data into the two groups.

The observed values, T and M, differ to some extent from one group to the other. If the observed values were not very different for subjects in Group 1 and Group 0, the two groups would be indistinguishable with respect to these variables. A multivariate test was done to investigate if these two variables were able to separate the two groups. In particular, a multivariate two-sample T^2-test was performed to determine if the mean value for Group 0, $(E(T_0), E(M_0))$, was significantly different from the mean value for Group 1, $(E(T_1), E(M_1))$.

The T^2-test examines whether the mean values of the tension and moment are equal for the two groups. The alternatives are:

$$(5) \qquad H_0 : \bar{\mu}_0 = \bar{\mu}_1 \quad vs \quad H_a : \bar{\mu}_0 \neq \bar{\mu}_1$$

Where $\bar{\mu}_i$ is the vector-valued expectation of the tension and moment for Group i = 0,1. The hypothesis test resulted in a p-value of 0.0001. Thus, there is evidence to conclude that the mean value of the vector (T, M) is significantly different for the two groups.

While this type of test is not strongly affected by lack of normality, a transformation of the variables T and M was applied to make the joint distribution of the transformed variables approximately normal. The Box-Cox, (1964), method was followed to transform the variables such that:

$$(6) \qquad (\bar{T}, \bar{M}) = \left(\frac{1}{\sqrt{T}}, \frac{1}{\sqrt{M}} \right)$$

The assumption of normality was reasonable and the T^2-test was performed using $(\tilde{T}, \tilde{M})$. The analysis of the transformed data supported the same conclusion reached above with the same significance level.

It is important to investigate if the moment measured in addition to the tension will significantly increase the separation of the two groups. A test of the hypothesis that the moment data is redundant in the presence of the tension data was conducted. This "additional information" test (see Appendix) resulted in a p-value of 0.78 implying that there is strong evidence to accept the hypothesis of moment being redundant. This suggests that the moment data does not add significant information for purposes of separating the serious-to-fatal and no-to-low injury groups. This conclusion leads us to ask if moment has any power to distinguish between the two groups.

Univariate tests were also done to compare the mean moment value for the two groups. A test was done to determine if the mean value of the moment is equal for the two groups. The alternatives are:

(7) $$H_0 : E(M)_0 = E(M)_1 \quad \text{vs} \quad H_a : E(M)_0 \neq E(M)_1$$

Where $E(M)_i$, is the expected value of the moment for Group $i = 0,1$. The necessary assumptions for this type of test (observations are independent, observations for each group are a sample from a population with a normal distribution and variances for the two independent groups are equal) seem to be reasonable. A univariate t-test to compare the two groups mean moment resulted in a p-value of 0.016. Thus, there is evidence to conclude that the mean moment for the two groups is not significantly different at level 0.01. The mean value of the moment can not be confidently used to separate the serious-to-fatal injury and no-to-low injury groups.

Finally, a univariate test to compare the mean tension value for the two groups was performed. As was done with moment, a test was done to determine whether the average value for the tension is equal for the two groups. The tension data is not likely to be normal (the distribution of Group 0 is skewed). Therefore, a Wilcoxon Rank Sum was used to compare the mean of the two groups. This test provides a non-parametric analog to the t-test used in the analysis of the mean moment. The only assumption is that the observations are independent. The p-value for this test is 0.0001. Thus, there is strong evidence to conclude that the average value of the tension for the two groups is significantly different. Tension can be applied to distinguish the no-to-low and serious-to-fatal injury groups.

Figure 12, a scatter plot of the tension and moment data with the no-to-low and serious-to-fatal injury cases indicated by legend, illustrates this point. The location of the means of the two groups is shown. The two means are significantly separated by tension. By contrast, the separation is not significant in terms of moment.

Paired Animal/Test Dummy Air Bag Deployment Tests for Neck Injury Potential
(Porcine Animal Subject and Three Year Old Test Dummy)
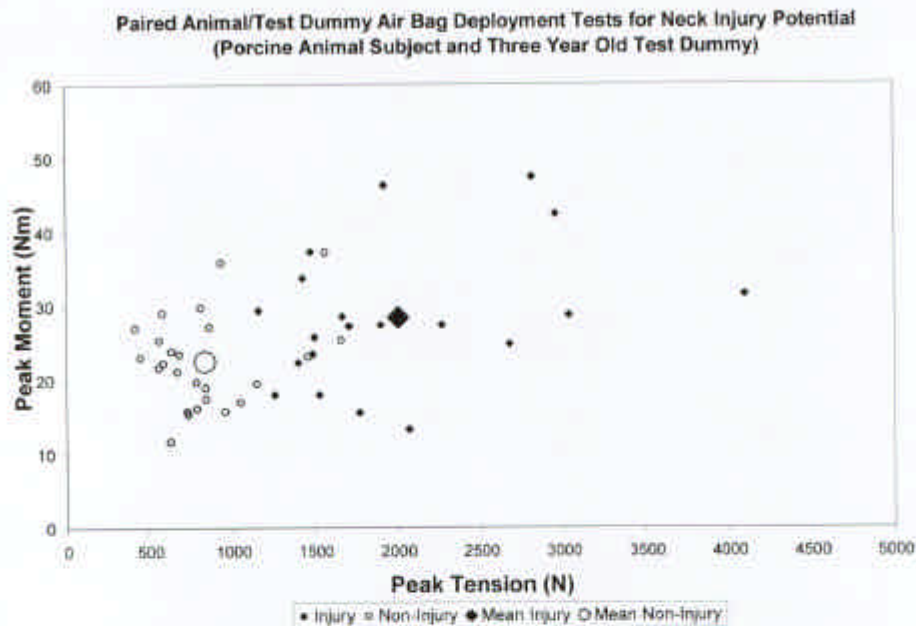

Figure 12. Scatter plot of biomechanical moment and tension data.

The additional information test gives evidence to conclude that the moment variable does not contribute anything significant beyond the information already available in tension for separating the serious-to-fatal injury and no-to-low injury groups. The univariate test conducted on moment reinforced this conclusion. The mean value of the moment is not significantly different for the two groups (the two mean moment values were not significantly different at 0.01 level). However, the mean value of the tension is significantly different for the two groups (the probability of observing a larger difference by chance alone is 0.0001).

There is confidence to conclude that, for all practical purposes, tension is the only valid variable. Moment is not informative in distinguishing between the serious-to-fatal injury and no-to-low injury groups. Therefore, if a linear combination of tension and moment is used to discriminate the two groups, the more weight given to moment will result in less power to discriminate the two groups. This conclusion can be supported more formally by trying to form a linear discriminant to classify observations of tension and moment into the two groups.

CLASSIFICATION ANALYSIS: ALLOCATION OF OBSERVATIONS TO GROUPS

In classification analysis a sampling subject whose membership is unknown is assigned to a group based on the parameter values associated with the subject. As in the previous statistical analysis of the observed measurements, all the available observations have been divided in two groups: Group 0 (no-to-low injury) and Group 1 (serious-to-fatal injury).

Three parameters were compared in their ability to classify the samples. Two separate approaches to linearly combining moment and tension were considered, N_{ij} and K (Mertz et al., 1997) as well as tension (T) alone. K is an instantaneous peak kernel of moment and tension defined as:

(8) $K = M_F + cF_T$

where c is a constant dependent on scale factors and characteristic dimensions of the neck (see Equation 1 as well). Fisher's discriminant function (Johnson and Wichern, 1998) was used to compare the performance of K, N_{ij} and T in classifying the samples. This technique transforms samples into a univariate space where the observations can be separated as much as possible. The discriminant function defines a threshold value in the univariate space that segregates the data into the separated groups.

The results of the classification procedure using Fisher's discriminant function for each of the parameters are summarized in Table 2. Each parameter was used to assign the samples to a group.

Table 2. CLASSIFICATIONS USING VARIOUS NECK INJURY ASSESSMENT PARAMETERS.

Parameter	Subject Group	Total Observations	Correct Classifications	Incorrect Classifications	Classification Rate
K	0	25	22	3	88%
	1	20	14	6	70%
N_{ij}	0	25	20	5	80%
	1	20	13	7	65%
T	0	25	22	3	88%
	1	20	16	4	80%

When the entire subject group is evaluated using the discriminant function, the largest percent correctly classified were observed when using tension alone.

One important way for judging the performance of any classification procedure is to calculate its "error rate" or misclassification probability. An estimate of the misclassification probabilities can be obtained using the cross validation method (Rencher, 1995) (sometimes referred as "hold –out" method). This approach does not depend on the parent population and, for moderately-sized sample (such as the one with this data set), it is nearly an unbiased estimate of the expected actual error rate, $E(AER)$. In this data set, where the parent population is unknown, the cross validation method gives the best available estimate of the probability that a classification function (K, N_{ij} and T) will misclassify a future observation based on the present sample.

Estimates of the probability of correct classifications are summarized in Table 3.

Table 3. ESTIMATES OF CORRECT INJURY CLASSIFICATIONS USING CROSS VALIDATION.

Classification Variable	T	K	N_{ij}
1-E(AER)	84.4%	80.0%	73.3%

The highest probability of correct classification rate results from using the tension variable. This result confirms what was observed in the previous two analyses: tension is the most informative variable in classifying an observation as either low-to-no injury (Group 0) or serious-to-fatal injury (Group 1). The multivariate and univariate statistical analyses suggested that there is evidence that moment has no power in classifying the observations. As expected, inclusion of moment degraded the predictive value of the analysis. Thus, the two injury assessment values that combine moment and tension, K and N_{ij}, have less power than the tension alone. Moment is only acting as a confounding variable.

CONSISTENT THRESHOLD ESTIMATE

Moment is not statistically significant in predicting neck injury, thus no risk curve can be estimated using these values. Moreover, K and N_{ij} have less power than the tension alone. The best estimated probability of injury is the one computed using tension alone. As it has been observed, tension is neither Normal nor Log-Normal distributed; its distribution function cannot be approximated by either of these two families. The form of the underlying distribution is unknown and there is no basis for a parametric risk analysis. Therefore, the best method for performing a risk analysis is a non-parametric approach. The Consistent Threshold Method (CTM) (Nusholtz and Mosier, 1999) will be used to determine the probability of injury. The CTM provides a non-parametric, maximum likelihood estimate.

There is not enough data collected in (Mertz et al., 1982b) and (Prasad and Daniel, 1984) to define strict threshold values for the risk of neck injury. Thus, based on the combined two data sets, the best that can be estimated are regions of low-risk, intermediate-risk, and high-risk. A low-risk region is a region below the first observed injury. A high-risk region is above the highest value for observed no-injury cases. The intermediate region is the one between the low and the high-risk regions.

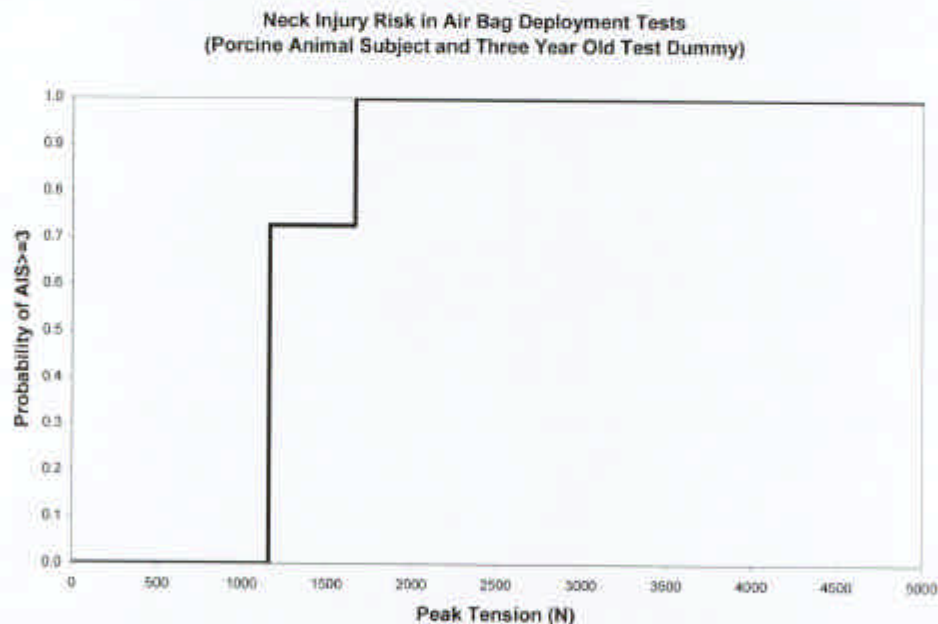The estimated risk of serious-to-fatal injury as a function of tension using the CTM is illustrated in Figure 13.

Figure 13. Estimated risk of serious-to-fatal injury based on tension data.

DISCUSSION

An analysis of the data collected in (Mertz et al., 1982b) and (Prasad and Daniel, 1984) calls into question the value of using moment, either on its own or in combination with tension, as an indicator of neck injury for air bag loading. This result may seem counterintuitive because moment can be a predictor of injury for inertial loading. The neck is bent under the action of air bag

loading. Bending a prismatic bar with homogeneous material properties will result in tensile and compressive stresses on the component being bent. Bending stress should be a contributor to ligament tissue failure in a similar manner to axial stress. This concept is the basis for the K and N_{TE} (the tension-extension part of the N_{ij} injury criterion). However, a rigorous analysis of the data shows a linear addition of extension moment to the tension data degrades the accuracy of injury prediction. Furthermore, injuries observed in the animal subjects are only related to tension.

Air bag deployment tests represent a complex set of loading conditions on the neck. The paired tests were conducted with the tacit assumption that the test device and the animal surrogate respond similarly to the application of this complex loading condition. Differences between the test device and the animal subject may offer a reason why the observed moment does not correlate with injury. The constant stress hypothesis that underlies the K and N_{ij} criteria assumes that the neck behaves as a prismatic, homogeneous bar. A biologic specimen has features that challenge this assumption.

The Hybrid III dummy response corridors were developed using inertial loading of the head while the torso is restrained by a seat belt and a seat back. A deploying air bag loads the head directly, creating a different type of load condition on the neck. Finally, inertial loading of the head causes the neck to behave as a cantilever beam with a point load at the free end. In simple terms, the maximum moment occurs when the deformation is at a maximum and deformation rate is at a minimum. By contrast, large moments can be imparted to the neck by air bag loading when the deformation rate is very high -- as much as an order of magnitude larger than what would be observed in inertial loading due to impact. Strain rate sensitivities in the test device may be influencing the moments recorded during the tests.

For the reasons stated above, the test device may not be appropriate for purposes of tests involving air bag loading when moment is considered. In addition it may call into question whether it is appropriate to assume that neck moment risk can be assessed using standard crash test dummies. Conversely, it may imply that moment is not an estimator of risk for air bag loading.

CONCLUSIONS

This study represents an analysis of the data contained in (Mertz et al., 1982b) and (Prasad and Daniel, 1984). Additional data of this type may change the numerical results but should not change the general conclusions.

The following conclusions apply only to air bag loading and may not apply to other types of loading.

Two analytical approaches were used to examine if the existing neck injury criteria for air bag loading are rigorously supported by the animal test data as good indicators of neck injury risk. A direct fit optimization analysis re-examined the constants used in the N_{ij} criterion. The results showed that no values of the constants used in to the N_{ij} criterion corresponding to compressive force and flexion moment will optimally separate the serious-to-fatal injury and no-to-low injury cases in the biomechanical data set. This result is expected: based on the experimental procedures employed in the original tests significant compression and flexion should not be generated. It is expected that only the tensile and extension intercepts should allow optimal identification of the injury risk.

The tension and moment data were examined to assess their injury prediction value using a variety of standard statistical techniques. Tension alone is the best predictor of injury. Although a combination of tension and moment can be used, all indications are that increasing the weight given to moment decreases the accuracy of the prediction. Moment is a confounding random

variable in the biomechanical data set: moment should not be used as an injury indicator it only decreases the accuracy.

Various neck injury criteria have been proposed which use tension and bending moment as separate variables and in linear combinations as predictors of injury for air bag loading. This investigation indicates that the use of moment is not supported by the biomechanical data used to form the basis for these proposed criteria. With the currently available test data and anthropomorphic test devices, only neck tension is a reliable predictor of injury potential for air bag loading.

ACKNOWLEDGEMENTS

REFERENCES

BOX, G. E. P., COX, D. R., (1964). "An Analysis of Transformations", Journal of the Royal Statistical Society (B), 26, No.2 , 211-252.

JOHNSON, R. A., WICHERN, D. W., (1998). "Applied Multivariate Statistical Analysis," Fourth Edition, Prentice Hall.

KIM, A., (1999). Personal Communication, November 1999.

MERTZ, H. J., (1999). Personal Communication, November 1999.

MERTZ, H. J., PRASAD, P., IRWIN, A. L., (1997). "Injury Risk Curves for Children and Adults in Frontal and Rear Collisions," SAE Paper 973318.

MERTZ, H. J., WEBER, D. A., (1982a). "Interpretations of the Impact Responses of a 3-Year-Old Child Dummy Relative to Child Injury Potential," Proceedings of the Ninth International Technical Conference on Experimental Safety Vehicles, SAE Paper 826048.

MERTZ, H. J., DRISCOLL, G. D., LENOX, J. B., NYQUIST, G. W., WEBER, D. A., (1982b). "Responses of Animals Exposed to Deployment of Various Passenger Inflatable Restraint System Concepts for a Variety of Collision Severities and Animal Positions," Proceedings of the Ninth International Technical Conference on Experimental Safety Vehicles, SAE 826047.

NHTSA, (2000a). "Final Rule, Occupant Crash Protection, Federal Motor Vehicle Safety Standard 208," National Highway Traffic Safety Administration, Department of Transportation.

NHTSA, (2000b). "Supplemental Notice of Proposed Rulemaking, Occupant Crash Protection, Federal Motor Vehicle Safety Standard 208," National Highway Traffic Safety Administration, Department of Transportation.

NHTSA, (1998). "Notice of Proposed Rulemaking, Occupant Crash Protection, Federal Motor Vehicle Safety Standard 208," National Highway Traffic Safety Administration, Department of Transportation.

NUSHOLTZ, G. S., SHI Y., XU, L., (2000). "Optimization of Single-point Frontal Airbag Fire Threshold," SAE Paper 2000-01-1009.

NUSHOLTZ, G. S., MOSIER, R., (1999). "Consistent Threshold Estimate for Doubly Censored Biomechanical Data," SAE Paper 1999-01-0714.

PRASAD, P., DANIEL, R., (1984). "A Biomechanical Analysis of Head, Neck, and Torso Injuries to Child Surrogates Due to Sudden Torso Acceleration," Proceedings of Twenty-Eighth Stapp Car Crash Conference, SAE Paper 841656.

RENCHER, A. C., (1995). "Methods of Multivariate Analysis," Wiley.

APPENDIX

Additional Information Test

The data under consideration consists of a random sample (T_{0j}, M_{0j}) with $j = 1,...,25$ (Group 0) that is assumed to be distributed $N_2(\overline{\mu}_0, \Sigma_0)$ and a second random sample (T_{1j}, M_{1j}) with $j = 1,...,20$ (Group 1) that is assumed to be distributed $N_2(\overline{\mu}_1, \Sigma_1)$. A further assumption is made that the two samples are independent. The T^2-statistic based on the full set of variables (both T and M) is given by:

$$T^2 = (\overline{y}_0 - \overline{y}_1)' \left(\left(\frac{1}{n_1} + \frac{1}{n_0} \right) S_{pl} \right)^{-1} (\overline{y}_0 - \overline{y}_1)$$

where S_{pl} is the pooled covariance matrix . $\overline{y}_i = \left(\sum_i \frac{T_{ij}}{n_i} \quad \sum_i \frac{M_{ij}}{n_i} \right)'$, $n_1 = 20$ and $n_0 = 25$.

The T^2-statistic based on the reduced set of variables (only Tension) is given by:

$$T_T^2 = \frac{20 \cdot 25}{45 \sigma_T^2} \left(\sum_i \frac{T_{0j}}{25} - \sum_i \frac{T_{1j}}{20} \right)^2$$

The hypothesis of M redundant is rejected at level α if:

$$F = 42 \left(\frac{T^2 - T_T^2}{43 + T_T^2} \right) \geq F_{(\alpha, 1, 42)}$$

For the combined data set (Mertz et al., 1982b, Prasad and Daniel, 1984), the value of the F-statistic above is 0.076, with a p-value equal to 0.78, implying that there is evidence to accept the hypothesis of M being redundant.

DISCUSSION

PAPER: **Dummy Neck Response During Airbag Loading and Injury Potential Assessment**

PRESENTER: *Guy Nusholtz, Daimler Chrysler*

QUESTION: *Frank Pintar, Medical College of Wisconsin*

This data always kind of baffles me because the way that the pigs were experimented with the airbags there's really no chin to produce tension in the pig, and yet the dummy data tells you that tension is the predictor. Would you care to comment on how those can be resolved?

ANSWER: I didn't go into a detailed look at the experimental setup. I do know you can, in fact, get some level of friction in the pigs. There should be some area in which the airbag can attach below the neck to produce tension, it's not completely, it doesn't completely eliminate all of it. First of all, the only thing you have to produce tension on the dummies, you don't have to produce it on the pigs in this case to get the tension. This is a correlator, it's a paired model.

Q: You correlate the tension?

A: You're also questioning the quality of the data that results. And what you're saying is okay, this showed that tension was an indicator of injury and moment is not. But even in that, you may not be able to extrapolate this to a live human because of the way of the experimental test setup. Someone would have to go through a detailed look at all of the different tests to see whether that part of it makes sense. I just took at face value what Bud and Priya had done in which they had made the claim that this was a reasonable representation of what would happen with a kid.

Q: It just seems to me that the actual injuries produced in the swine were probably more due to bending just because of the nature of the subject, the swine. If you look at some of those old pictures there's basically a smooth surface that the airbag hits, there's no protruding chin like in the dummy. So it would seem to me that the high cervical injuries that they were getting in the swine had a lot more to do with the bending of the head back. And that's why it is interesting that when you put a dummy in there you get tension as a predictor?

A: I really can't speak for Bud, but for a short moment I will. When he explained it to me he said that the injuries that they saw looked like tension injuries. The way the tears and the damage looked like you had pulled on both sides of the neck and you had damage on both sides, which would indicate -- it looked like to him that it was a tension-type injury.

Q: *Erik Takhounts, NHTSA*

The dummy's neck, I think, the way I recall it, has some extra moment added to it and the argument was that this should represent actual musculature. Whereas pigs, obviously they didn't have that. Can you comment on this? Would it introduce some extra variable in your case, in your analysis?

A: Well, the extra musculature in the dummy is going to add a force on the head, but it doesn't represent B in a person it is not going to represent the force on the condoyles or where the injury is. So what it is going to be doing is you are going to be using this muscle as an estimator of force on the condoyles. That may not be accurate. In other words, the actual moments at the condoyles might be very different and, in fact, that might indicate that it is very different than what you're actually measuring in the dummy.

Q: So, you're saying that we'll have a dummy with low moment in the occipital condoyles but has springs that are representing muscles, they may better correlate with injury?

A: That may better correlate, yes.

Q: *Roger Nightingale, Duke University*

I wanted to comment on what looks like a tension injury and what doesn't. We're producing AO dislocations in cadavers, human cadavers right now with pure extension and flexion bending moments. So a pure moment can produce an AO dislocation without any fracture. So, I don't know that you can necessarily look at these injuries and say that's a tension injury.

The other question I had was you're using inverted dynamics data from the pigs, or not you, Priya and Bud were. And given the difficulties we just saw that the Medical College of Wisconsin is having calculating moments and tension at the condyles based on -- was it accelerometer and video data I think they're using, how confident are you in those numbers, basically, the pig numbers?

A: Once again, you're asking me to speak for Bud and Priya. They ran the experiments. They were confident that the numbers represented something. Whether that's true or not, I can't really address those particular type of answers. But I would agree that's a justifiable question to ask, what is the confidence that those actually represent something real. In defense of it, it was an attempt to try and get a handle and put a numerical value on this. These were run a long time ago and they are sort of now used as the basis for assessment criteria. So, if we just take them as they are and go through a process these are the results of what comes out of those numbers, good or bad.